View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **135**, 545–567 (1988)

Estimation of Cauchy Data for a First-Order Nonlinear Hyperbolic Equation

KEVIN A. GRASSE AND L. W. WHITE

Department of Mathematics, University of Oklahoma, Norman, Oklahoma 73019

Submitted by V. Lakshmikantham

Received February 4, 1985

1. INTRODUCTION

Let Ω be an open interval in $\mathbb{R}$. We study the Cauchy problem

$$\begin{aligned} u_t + uu_x + au &= \phi(x, t) \qquad \text{in} \quad \Omega \times (0, t) \\ u(x, 0) &= u_0(x). \end{aligned} \tag{1.1}$$

Equation (1.1) may be viewed as a model of a one-dimensional flow in which the constant $a > 0$ represents a friction term, ϕ a forcing term related to a pressure gradient, $u_0(x)$ the initial velocity at point x, and $u(x, t)$ the velocity of fluid at the point x and at time t [5]. It is well known (cf. [11]) that (1.1) is associated with an initial value problem for the system of ordinary differential equations

$$\begin{aligned} \frac{dx}{dt} &= z, & x(0, \xi) &= \xi \\ \frac{dz}{dt} &= -az + \phi(x, t), & z(0, \xi) &= u_0(\xi) \end{aligned} \tag{1.2}$$

called the characteristic equations for (1.1).

We study the following problems for (1.1). Suppose that M observations $\{z_k\}_{k=1}^M$ at positions $\{x_k\}_{k=1}^M$ are given at time t_0. We wish to estimate the initial condition u_0 from among an admissible set Q_{ad} of initial conditions. Roughly, our approach is to first solve the characteristic equations with the data as initial conditions backward to $t = 0$. That is, solve

$$\begin{aligned} \frac{dx}{dt} &= z, & x(t_0) &= x_k \\ \frac{dz}{dt} &= -az + \phi(x, t), & z(t_0) &= z_k \end{aligned} \tag{1.3}$$

0022-247X/88 $3.00

Copyright © 1988 by Academic Press, Inc.
All rights of reproduction in any form reserved.

to find $x(0; x_k, z_k) = \xi_k$ and $z(0; x_k, z_k) = \zeta_k$ for $k = 1, ..., M$, where $(x(t; x_k, z_k),\ z(t; x_k, z_k))$ denotes the solution of (1.3). Having thus obtained $\{\xi_k\}_{k=1}^M$ and $\{\zeta_k\}_{k=1}^M$, we introduce a fit-to-data functional given by

$$J(u_0) = \sum_{k=1}^{M} (u_0(\xi_k) - \zeta_k)^2 \tag{1.4}$$

and consider the problem

$$\begin{aligned} &\text{Find } \hat{u}_0 \in Q_{\text{ad}} \text{ such that} \\ &J(\hat{u}_0) = \min\{J(u_0) \colon u_0 \in Q_{\text{ad}}\}. \end{aligned} \tag{1.5}$$

In the following sections we make a detailed study of the above program. In particular in Section 2 we examine properties of solutions of (1.1) in order to specify requirements needed for Q_{ad}. In Section 3 we formulate precisely problem (1.5), obtain identifiability results, and give a useful approximation theory for the estimated initial condition using cubic splines. Finally, in Section 4 we present the results of numerical experiments.

A well-known feature associated with the problem (1.1) is that solutions may exhibit shocks under certain conditions [11]. Shocks may occur in a time interval $(0, t)$ for (1.1) when the Cauchy data are such that characteristic curves associated with (1.2) for different ξ intersect. In such a case the classical C^1 solution of (1.1) does not exist in $\Omega \times (0, T)$, however, a generalized entropy solution may. A study of estimation problems for (1.1) in the presence of possible shocks is contained in [12]. Here we provide compatibility conditions depending on u_0, a, ϕ, and T to guarantee the existence of classical solutions. These conditions are incorporated into the admissibility constraints in specifying Q_{ad}. Additional constraints upon Q_{ad} to guarantee compactness are necessary for existence of a solution to (1.5).

In addition to obtaining existence and approximation results, we also treat identifiability. We show that (i) there is a continuous dependence of the solution of (1.5) upon the data, (ii) although the solution is not unique, the set of solutions of (1.5) is closed and convex, and (iii) the diameter of the solution set may be estimated from the spacing of the data.

Our results here seem to have application to the data assimilation problem in meteorology [8]. The data assimilation problem may be stated as follows:

> given a (computer) model that is generating a solution to a fluid flow problem and given data obtained from measurements of fluid velocities at a specified time, find a method to update the model by incorporating these observations to produce a solution that is close to reality but that does not introduce instabilities.

The problem studied here is in essentially a least-squares regression of data. Here, however, the mathematical model is used to carry observations to $t=0$. From this transforming data an initial condition is estimated from a constraint set. Our method produces a smooth estimate of initial condition that produces a solution close to observations but that does not undergo a shock. This initial condition may then be used to solve the model equations to provide an update that takes into account observations.

2. Properties of Solutions of the Model Equations

Since we are interested in classical solutions of (1.1), we begin by giving the families of systems of characteristic ordinary differential equations associated with (1.1). Thus, we have

$$\begin{aligned} \frac{dx}{dt} &= z, & x(\xi, t_0) &= \xi \\ \frac{dz}{dt} &= -az + \phi(x, t), & z(\xi, t_0) &= u_0(\xi). \end{aligned} \tag{2.1}$$

The estimation scheme discussed in the following sections requires that the solution of (1.1) depend continuously (in some appropriate topology) on the parameter u_0. We are concerned primarily with classical solutions of (1.1) constructed from solution of (2.1). Consequently, the continuous dependence of the solution on parameters may be obtained from standard theorems in the theory of ordinary differential equations. One restrictive feature of this approach is that it is necessary to impose certain conditions on the parameters. These conditions are needed to insure the existence of the solution of characteristic ordinary differential equations in a specified time interval and the capability to invert a portion of the solution mapping, which is required in the construction of the solution of the Cauchy problem, We will be more explicit about the nature of these conditions later.

In the remainder of this section we briefly review the method of characteristics and cite appropriate general theorems that enable us to construct the solution of (1.1) from the characteristic ordinary differential equations. At the same time we show the continuous dependence of the solution of the Cauchy problem on the parameters.

The existence of solutions of (2.1) is a consequence of the following standard theorem.

Theorem 2.1. *Let $I \subseteq \mathbb{R}$ be an interval and let $V \subseteq \mathbb{R}^n$ be an open set. Let*

$$G: I \times V \to \mathbb{R}^n$$

be a mapping such that

(i) *G is continuous,*

(ii) *for each $(t_0, y_0) \in I \times V$ there exists an open neighborhood $I_0 \times V_0$ of (t_0, y_0) in $I \times V$ and a real constant K such that*

$$\|G(t, y_1) - G(t, y_2)\| \leqslant K \|y_1 - y_2\|$$

for every $t \in I_0$ and y_1, y_2 in V_0.

Then for each $(s, \xi) \in I \times V$ there exists a unique solution

$$t \mapsto \mu(t, s, \xi) \tag{2.2}$$

of the initial value problem

$$\begin{aligned} \dot{y} &= G(t, y) \\ y(s) &= \xi. \end{aligned} \tag{2.3}$$

The solution (2.2) *is defined on a maximal* (*but possibly proper*) *open subinterval of I containing s. Furthermore, the mapping μ is defined on an open subset of $I \times I \times V$ and is continuous in all three variables.*

Proof. See [3].

COROLLARY 2.1. *Suppose that in addition to* (i) *and* (ii) *the mapping G satisfies*

(iii) *$(t, y) \mapsto G(t, y)$ of $I \times V$ into $\mathbb{R}^n$ is C^1 and the partials $D_1 G$ and $D_2 G$ are continuous on $I \times V$.*

Then the solution mapping μ is C^1 and the partial derivatives $D_1 \mu$, $D_2 \mu$, $D_3 \mu$ are continuous on the domain of definition of μ in $I \times I \times V$.

Proof. See [3].

Remarks 2.1. (a) Condition (ii) of Theorem 2.1 can be described as a local Lipschitz condition in V locally uniform in I.

(b) The solution mapping μ is sometimes called the flow of the ordinary differential equation $\dot{y} = G(t, y)$.

In order to apply these results to the system of characteristic ordinary differential equations (2.1), we set $y = (x, z)$ and

$$G(t, y) = (z, -az + \phi(x, t)).$$

We assume

$$a \in (0, \infty) \text{ is constant;} \tag{H1}$$

$$\phi \in C^1(\mathbb{R} \times [0, \infty), \mathbb{R}) \text{ with}$$

$$|\phi(x, t)| + |\phi_x(x, t)| \leqslant \beta \tag{H2}$$

for all (x, t) in $\mathbb{R} \times [0, \infty)$.

We view G as a mapping

$$G\colon [0, \infty) \times \mathbb{R}^2 \to \mathbb{R}^2.$$

It is clear that G is C^1 from $[0, \infty) \times \mathbb{R}^2$ to $\mathbb{R}^2$. Furthermore, G satisfies a Lipschitz condition in $\mathbb{R}^2$ that is uniform in $[0, \infty)$ with constant $K = 1 + a + \beta$. This follows easily from the mean-value theorem.

Remark 2.2. The following estimate is a consequence of Gronwall's inquality and is standard (see [3]):

$$\|\mu(t, s, \xi_1) - \mu(t, s, \xi_2)\| \leqslant \|\xi_1 - \xi_2\| \, e^{K|t-s|},$$

where K is a global Lipschitz constant $K = 1 + a + \beta$.

We now have the following theorem.

THEOREM 2.2. *Let* (H1) *and* (H2) *hold. There exists a unique solution*

$$t \mapsto (X(t, \xi, \eta), Z(t, \xi, \eta)) \tag{2.4}$$

of the initial value problem

$$\begin{aligned} \frac{dx}{dt} &= z, & x(0) &= \xi \\ \frac{dz}{dt} &= -az + \phi(x, t), & z(0) &= \eta \end{aligned} \tag{2.1$'$}$$

that is defined for all $t \geqslant 0$. The solution functions X and Z are defined and continuous on an open subset of $[0, \infty) \times \mathbb{R}^2$ and they have first partial derivatives with respect to t, ξ, η which are continuous on the domain of definition of X and Z.

Proof. Under the assumption (H2) it follows that the right side of the characteristic ordinary differential equations (2.1)$'$ satisfies a global Lipschitz condition in the variables x and z and uniformly for $t \geqslant 0$. The fact that (2.4) exists for each $t \geqslant 0$ follows from a standard extension theorem in the theory of ordinary differential equations [2].

We briefly recall how a classical soluttion $u(x, t)$ of problem (1.1) is obtained from the solution (2.4) of the characteristic equations (2.1). We set

$$x = X(t, \xi, u_0(\xi)) \tag{2.5}$$

and attempt to solve this equation for ξ in terms of x and t, say

$$\xi = W(t, x)$$

(suppressing for the moment u_0). If this can be done, then

$$u(x, t) = Z(t, W(t, x), u_0(W(t, x)))$$

is the desired solution of (1.1). However, one cannot take for granted that the solution is defined for large values of t. It is true that $u(x, t)$ is defined on some open subset of $\mathbb{R} \times [0, \infty)$ containing the x-axis

$$\{(x, 0) : x \in \mathbb{R}\}$$

since $X(0, \xi, u_0(\xi)) = \xi$ for every $\xi \in \mathbb{R}$ so that

$$\frac{\partial}{\partial \xi} X(0, \xi, u_0(\xi)) = 1 \neq 0.$$

Hence, the inverse function theorem allows us to invert (2.5) at points near the x-axis. Other assumptions in addition to (H2) are required to insure that (2.5) can be inverted for "large" values of t.

For example, suppose we wish to guarantee that $u(x, t)$ is defined for every $(x, t) \in \mathbb{R} \times [0, T]$, where $T > 0$ is specified in advance. In view of the above discussion, it is required that the function

$$\xi \mapsto X(t, \xi, u_0(\xi))$$

be a C^1 diffeomorphism of $\mathbb{R}$ into $\mathbb{R}$ for each $t \in [0, T]$. A sufficient condition that this occur is that for each $t \in [0, T]$ there exist an $\alpha_t > 0$ such that

$$\frac{\partial}{\partial \xi} X(t, \xi, u_0(\xi)) \geqslant \alpha_t \qquad \text{for every } \xi \in \mathbb{R}. \tag{2.6}$$

Here we assume that u_0 is a C^1 function so that the derivative makes sense. Hence, we are reduced to formulating conditions on u_0 and T (as well as a and ϕ) which imply (2.6) for $0 \leqslant t \leqslant T$. At this point some notation is useful.

Notation 2.1. Let

$$U=\{u_0\in C^1(\mathbb{R},\mathbb{R})\colon \sup\{|u_0(x)|+|u_0'(x)|\colon x\in\mathbb{R}\}<\infty\}.$$

The U is a vector subspace of $C^1(\mathbb{R},\mathbb{R})$ and U is a Banach space under the norm

$$\|u_0\|=\sup\{|u_0(x)|+|u_0'(x)|\colon x\in\mathbb{R}\}.$$

Given that $(X(t,\xi,\eta), Z(t,\xi,\eta))$ denotes the solution of (2.1)′, it is clear from the second equation of (2.1)′ that X and Z are related by the equation

$$Z(t,\xi,\eta)=e^{-at}\eta+e^{-at}\int_0^t e^{as}\,\phi(X(s,\xi,\eta),s)\,ds. \tag{2.7}$$

From the first equation of (2.1)′ and (2.7), we obtain

$$X(t,\xi,\eta)=\xi+\frac{1}{a}(1-e^{-at})\,\eta+\int_0^t\frac{1}{a}(1-e^{-a(t-s)})\,\phi(X(s,\xi,\eta),s)\,ds. \tag{2.8}$$

We assume that (H1) and (H2) are satisfied so that $X(t,\xi,\eta)$ and $Z(t,\xi,\eta)$ are defined for every $t\geqslant 0$. In the initial condition in (1.1) we take $u_0\in U$. These assumptions imply that the derivatives $\|D\phi(x,t)\|$ and $|u'(x)|$ are globally bounded. This discussion leads us to consider the derivative $(\partial/\partial\xi)\,X(t,\xi,u_0(\xi))$ and, setting $\eta=u_0(\xi)$ in (2.8), we obtain

$$\begin{aligned}\frac{\partial}{\partial\xi}X(t,\xi,u_0(\xi))&=1+\frac{1}{a}(i-e^{-at})\,u_0'(\xi)\\&\quad+\frac{1}{a}\int_0^t(1-e^{-a(t-s)})\,\phi_x(X(s,\xi,u_0(\xi)),s)\,\frac{\partial}{\partial\xi}X(s,\xi,u_0(\xi))\,ds.\end{aligned} \tag{2.9}$$

The estimate $(1/a)(1-e^{-at})\leqslant 1/a$ for $t\geqslant 0$ and the Gronwall inequality applied to (2.9) yield

$$\left|\frac{\partial}{\partial\xi}X(t,\xi,u_0(\xi))\right|\leqslant\left(1+\frac{1}{a}|u_0'|\right)\exp(|\phi_x|\,t/a)\qquad\text{for}\quad t\geqslant 0, \tag{2.10}$$

where

$$|u_0'|=\sup\{|u_0'(\xi)|\colon \xi\in\mathbb{R}\} \tag{2.11}$$

and

$$|\phi_x|=\sup\{|\phi_x(\xi,t)|\colon(\xi,t)\in\mathbb{R}\times[0,\infty)\}. \tag{2.12}$$

In view of (2.7), we see that (2.6) is satisfied for some $\alpha_t > 0$ if

$$\sup_{\xi \in \mathbb{R}} \left| \frac{1}{a}(1-e^{-at})\, u_0'(\xi) + \int_0^t \frac{1}{a}(1-e^{-a(t-s)})\, \phi_x(X(s, \xi, u_0(\xi))) \times \frac{\partial}{\partial \xi} X(s, \xi, u_0(\xi))\, ds \right| < 1 \tag{2.13}$$

From (2.10) we obtain

$$\begin{aligned} &\left| \frac{1}{a}(1-e^{-at})\, u_0'(\xi) + \int_0^t \frac{1}{a}(1-e^{-a(t-s)})\, \phi_x(X(s, \xi, u_0(\xi)), s) \frac{\partial}{\partial \xi} X(s, \xi, u_0(\xi))\, ds \right| \\ &\leqslant \frac{|u_0'|}{a} + \frac{|\phi_x|}{a}\left(1 + \frac{|u_0'|}{a}\right) \int_0^t \exp(|\phi_x|\, s/a)\, ds \\ &= \frac{|u_0'|}{a} + \left(1 + \frac{|u_0'|}{a}\right)(\exp(|\phi_x|\, t/a) - 1) \\ &= -1 + \left(1 + \frac{|u_0'|}{a}\right) \exp(|\phi_x|\, t/a). \end{aligned}$$

Thus, in order for (2.13) to be satisfied it suffices that

$$-1 + (1 + |u_0'|/a) \exp(|\phi_x|\, t/a) < 1$$

or

$$(1 + |u_0'|/a) \exp(|\phi_x|\, t/a) < 2. \tag{2.14}$$

We summarize the above discussion in the following theorem.

THEOREM 2.3. *Consider problem* (1.1), *let* (H1) *and* (H2) *hold, let* $u_0 \in U$, *and let* $|u_0'|$ *and* $|\phi_x|$ *be defined by* (2.11) *and* (2.12), *respectively. Then* (1.1) *has a unique* C^1 *solution* $u(x, t)$ *defined for* $(x, t) \in \mathbb{R} \times [0, T]$ *provided that* $|u_0'| < a$ *and*

$$(1 + |u_0'|) \exp(|\phi_x|\, T/a) < 2.$$

Our next objective is to examine the dependence of the solution $u(x, t)$ of (1.1) upon u_0.

LEMMA 2.1. *For the solution* (2.4) *of the initial value problem* (2.1)′, *the mapping*

$$(t, \xi, u_0) \mapsto (X(t, \xi, u_0(\xi)), Z(t, \xi, u_0(\xi)))$$

of $[0, \infty) \times \mathbb{R} \times U$ *into* $\mathbb{R}^2$ *is* C^1 *in the variables* t, ξ, u_0.

Proof. This is an immediate consequence of Theorem 2.3 and the fact that the evaluation mapping of $U \times \mathbb{R}$ into $\mathbb{R}$ given by $(u_0, \xi) \mapsto u_0(\xi)$ is C^1.

Notation 2.2. Define the subset U_α by

$$U_\alpha = \{u_0 \in U: \sup\{|u_0'(x)|: x \in \mathbb{R}\} < \alpha\}.$$

Certainly, U_α is open in U.

THEOREM 2.4. *Fix* $a > 0$, $T > 0$ *and choose* $\alpha > 0$ *so that* $\alpha < a$ *and*

$$\left(1 + \frac{\alpha}{a}\right) \exp\left(\frac{\beta T}{a}\right) < 2. \tag{2.15}$$

Then for each $u_0 \in U_\alpha$ *the problem* (1.1) *has a solution* $u(x, t; u_0)$ *defined for every* $(x, t) \in \mathbb{R} \times [0, T]$. *Furthermore, the mapping*

$$(x, t, u_0) \mapsto u(x, t; u_0)$$

of $\mathbb{R} \times [0, T] \times U_\alpha$ *into* $\mathbb{R}$ *is* C^1 *in all the variables* x, t, u_0.

Proof. The conditions $u_0 \in U_\alpha$ and (2.15) yield the inequality (2.14) so the existence of $u(x, t; u_0)$ on $\mathbb{R} \times [0, T]$ is a consequence of Theorem 2.3. Consider the solution (2.4) of (2.1)′. By Lemma 2.1 the mappings $X(t, \xi, u_0(\xi))$ and $Z(t, \xi, u_0(\xi))$ are C^1 in the variables t, ξ, u_0. The inequality (2.15) guarantees that

$$\frac{\partial}{\partial \xi} X(t, \xi, u_0(\xi)) \neq 0$$

for $0 \leqslant t \leqslant T$ and $\xi \in \mathbb{R}$ (cf. the discussion in the proof of Theorem 2.3), so the implicit function theorem implies that we can solve the equation

$$x = X(t, \xi, u_0(\xi))$$

for ξ as a function of t, x, u_0. More precisely there exists a C^1 mapping

$$W: [0, T] \times \mathbb{R} \times U_\alpha \to \mathbb{R}$$

such that $x = X(t, \xi, u_0(\xi))$ if and only if $\xi = W(t, x, u_0)$. The solution $u(x, t; u_0)$ is given by the formula

$$u(x, t; u_0) = Z(t, W(t, x, u_0), u_0(W(t, x, u_0))).$$

This exhibits the mapping $(x, t, u_0) \mapsto u(x, t; u_0)$ as a composition of C^1 mappings and thus proves that $u(x, t; u_0)$ is a C^1 function of its variables.

As a particular consequence $u(x, t; u_0)$ is uniformly continuous on compact subsets of its domain $\mathbb{R} \times [0, T] \times U_\alpha$. This immediately yields the following useful corollary.

COROLLARY 2.2. *Let $(u_0^n) \subseteq U_\alpha$ be a sequence such that $u_0^n \to u_0 \in U_\alpha$, the convergence being with respect to the norm on U. Then*

$$u(x, t; u_0^n) \to u(x, t; u_0)$$

uniformly on compact subsets of $\mathbb{R} \times [0, T]$.

3. THE IDENTIFICATION PROBLEM

We begin by more precisely posing the problem to estimate Cauchy data of (1.1). Suppose we are given M observations $\{z_k\}_{k=1}^M$ at locations $\{x_k\}_{k=1}^M$ and at time $t = t_0 > 0$. We assume there are positive numbers X and T such that for $1 \leqslant k \leqslant M$,

$$\begin{aligned} &0 \leqslant x_k \leqslant X \\ &0 < t_0 \leqslant T. \end{aligned} \tag{3.1}$$

For ease we assume $x_1 < x_2 < \cdots < x_M$. Given these data we consider M characteristic ordinary differential equations with initial conditions determined by the observed data. That is, we have for $k = 1, ..., M$ that

$$\begin{aligned} \frac{dx}{dt} &= z, & x(t_0) &= x_k \\ \frac{dz}{dt} &= -az + \phi(x, t), & z(t_0) &= z_k. \end{aligned} \tag{3.2}$$

We denote the solution of (3.2) associated with the kth observation by

$$t \mapsto (x_k(t), z_k(t)).$$

When $t=0$, we set $\xi_k = x_k(0)$ and $\zeta_k = z_k(0)$. We make the following assumption that is motivated by Theorem 2.3:

$$\phi,\ T, \text{ and } a \text{ satisfy } \exp(|\phi_x|\, T/a) < 2. \tag{H3}$$

The condition (H3) may be viewed as a compatibility condition among the parameters ϕ, T, and a that enables us to formulate problems with classical solutions. Assumption (H3) is important in specifying the set of admissible Cauchy data.

Let $M_D = \max\{\|(x_i, z_i) - (x_j, z_j)\|: i, j = 1, ..., M\}$. Then with Lipschitz constant $1 + a + \beta$ we see from Remark 2.1(b) that for any $i, j = 1, ..., M$,

$$|\xi_i - \xi_j| \leqslant \|(\xi_i - \xi_j, \zeta_i - \zeta_j)\| \leqslant M_D \exp((1 + a + \beta)\, T) = k.$$

Hence, setting $\xi_D = \min\{\xi_i \text{ for } i = 1, ..., N \text{ and } 0\}$ we define Ω to be a bounded open interval such that

$$[\xi_D, \xi_D + k] \subset \Omega$$

and note that $[0, X] \subset \Omega$.

We consider functions with compact support contained in Ω and when necessary we may extend the function outside of Ω by zero. In particular, set $Q = C_0^1(\Omega)$. Since $H_0^2(\Omega)$ embeds compactly into Q [1], we consider primarily $H_0^2(\Omega)$.

Let us set $\gamma = a/(\text{meas}\ \Omega)^{1/2}$ and consider the inequality

$$\|u_0\|_{H_0^2(\Omega)} \leqslant \gamma(2 \exp(-|\phi_x|\, T/a) - 1). \tag{3.3}$$

If (3.3) holds, then by integration of u_0 it follows that condition (2.14) holds. Hence, by Theorem 2.3 there is a unique C^1 solution in $\Omega \times [0, T]$ (in fact in $\mathbb{R} \times [0, T]$ if we extend u_0 by zero outside of Ω).

To proceed we specify an appropriate optimization problem. We choose to compare parameters by means of a fit-to-data functional given by

$$J(u_0) = \sum_{k=1}^{M} (\zeta_k - u_0(\xi_k))^2. \tag{3.4}$$

Moreover, we consider the problem with the set of admissible parameters given by

$$Q_{\text{ad}} = \{u_0 \in H_0^2(\Omega): u_0 \text{ satisfies } (3.3)\}. \tag{3.5}$$

Remark 3.1. From the discussion above it follows that for each $u_0 \in Q_{\text{ad}}$ there exists a classical solution in $\Omega \times [0, T]$. Moreover, since Q_{ad} is a closed ball in $H_0^2(\Omega)$, it is weakly closed and weakly compact.

The optimization prroblem associated with the estimation problem is given by

$$\text{Find } u_0^* \in Q_{\text{ad}} \text{ such that } J(u_0^*) = \inf\{J(u_0) \colon u_0 \in Q_{\text{ad}}\}. \tag{3.6}$$

THEOREM 3.1. *Suppose* (H1)–(H3) *hold. There exists a solution to problem* (3.6).

Proof. Since $Q_{\text{ad}} \neq \phi$ let $\{u_0^n\}_{n=1}^{\infty}$ be a minimizing sequence. Thus, we have

$$J(u_0^n) \to d = \inf_{u_0 \in Q_{\text{ad}}} J(u_0).$$

From Remark 3.1 it follows that there is a subsequence $\{u_0^{n_i}\}_{i=1}^{\infty}$ such that $\{u_0^{n_i}\}_{i=1}^{\infty}$ converges weakly in $H_0^2(\Omega)$. Denote the limit by u_0^* and note that $u_0^* \in Q_{\text{ad}}$. Since the embedding of $H_0^2(\Omega)$ into $C^1(\bar{\Omega})$ is compact, it follows that the subsequence may be chosen such that

$$u_0^{n_i} \to u_0^*$$

in $C_0^1(\bar{\Omega})$. Accordingly, we see that

$$u_0^{n_i}(\xi_k) \to u_0^*(\xi_k)$$

and

$$J(u_0^{n_i} \to J(u_0^*)$$

as $i \to \infty$. Therefore, we have $J(u_0^*) = d$ and u_0^* is a solution.

Remark 3.2. The set Q_{ad} given above is an example of one possible such set. Of course, other admissible sets may be used. However, such sets should be contained in $C_0^1(\Omega)$ and its members satisfy an inequality such as that in Theorem 2.3. If Q_{ad} is specified as a bounded closed subset of a finite-dimensional subspace of $C_0^1(\bar{\Omega})$, we automatically have compactness.

In the case that Q_{ad} is infinite dimensional it is necessary to approximate (3.6) by a sequence of finite-dimensional problems. This amounts to considering (3.6) with Q_{ad} replaced by a set Q_{ad}^N in a finite-dimensional subspace. These finite-dimensional subsets should approximate Q_{ad} in some manner. We now demonstrate this for the Q_{ad} given in (3.5).

Let Ω be partitioned into M equal subintervals and consider cubic B-splines defined on the resulting mesh; see [9, 10]. We define a subspace $S^M(\Omega) = \text{span}\{B_k\}_{k=1}^{l(M)}$, where $\{B_k\}_{k=1}^{l(M)}$ is a basis composed of cubic spline functions such that $S^M(\Omega) \subset H_0^2(\Omega)$; see [9, 10].

We use the notation

$$\|\varphi\|_{H_0^2(\Omega)} = \left(\int_\Omega \varphi_{xx}^2 \, dx \right)^{1/2}$$

and observe that $\|\cdot\|_{H_0^2(\Omega)}$ is a norm on $H_0^2(\Omega)$. Denote the interpolation operator by I^M,

$$I^M : H_0^2(\Omega) \to S^M(\Omega);$$

see [10]. We use the property

$$\|I^M \varphi\|_{H_0^2(\Omega)} \leqslant \|\varphi\|_{H_0^2(\Omega)} \tag{3.7}$$

for any $\varphi \in H_0^2(\Omega)$, [10].

Lemma 3.1. *If $\varphi \in H_0^2(\Omega)$, then $I^M \varphi \to \varphi$ strongly in $H_0^2(\Omega)$.*

Proof. If $\varphi \in H_0^2(\Omega)$, then by (3.7)

$$\|I^M \varphi\|_{H_0^2(\Omega)} \leqslant \|\varphi\|_{H_0^2(\Omega)}.$$

Moreover, if $\psi \in C_0^\infty(\Omega)$, then

$$\|I^M \psi - \psi\|_{H_0^2(\Omega)} \leqslant C \frac{1}{M^2} \|D^4 \psi\|_{L^2(\Omega)}.$$

Therefore since $C_0^\infty(\Omega)$ is dense in $H_0^2(\Omega)$ we have that for any $\varphi \in H_0^2(\Omega)$ there exists $\psi \in C_0^\infty(\Omega)$ such that given $\varepsilon > 0$ then $\|\varphi - \psi\|_{H_0^2(\Omega)} \leqslant \varepsilon/3$. Hence we see that

$$\|I^M \varphi - \varphi\|_{H_0^2(\Omega)} \leqslant \|I^M(\varphi - \psi)\|_{H_0^2(\Omega)} + \|I^M \psi - \psi\|_{H_0^2(\Omega)} + \|\psi - \varphi\|_{H_0^2(\Omega)},$$

From inequality (3.7) it follows that

$$\|I^M \varphi - \varphi\|_{H_0^2(\Omega)} \leqslant \frac{2\varepsilon}{3} + \frac{C}{M^2} \|D^4 \psi\|_{L^2(\Omega)}.$$

There exists $M(\varepsilon) \in \mathbb{N}$ such that if $M \geqslant M(\varepsilon)$ then

$$\frac{C}{M^2} \|D^4 \psi\|_{L^2(\Omega)} < \frac{\varepsilon}{3}.$$

Thus, we see that for any $M \geqslant M(\varepsilon)$,

$$\|I^M \varphi - \varphi\|_{H_0^2(\Omega)} \leqslant \varepsilon$$

and $I^M \varphi \to \varphi$ in $H_0^2(\Omega)$ as $M \to \infty$.

We define Q_{ad}^{M} by

$$Q_{\mathrm{ad}}^{M} = S^{M}(\Omega) \cap Q_{\mathrm{ad}}$$

for $M \in \mathbb{N}$. Hence, for Q_{ad} defined in (3.5) we have immediately the following.

LEMMA 3.2. *Let Q_{ad} be given by* (3.5). *If Q_{ad}^{M} is defined by* (3.8), *then for any $u_0 \in Q_{\mathrm{ad}}$ we have $I^{M} u_0 \in Q_{\mathrm{ad}}^{M}$.*

The finite-dimensional problem corresponding to (3.6) is

$$\text{Find } u_{0_M}^{*} \in Q_{\mathrm{ad}}^{M} \text{ such that } J(u_{0_M}^{*}) = \inf\{J(u_0) \colon u_0 \in Q_{\mathrm{ad}}^{M}\}. \tag{3.9}$$

The existence of a solution of (3.9) is proved in a manner similar to the proof of Theorem 3.1.

THEOREM 3.2. *For each $M \in \mathbb{N}$ there exists a solution $u_{0_M}^{*}$ to problem* (3.9).

We now consider the limiting behavior of the sequence $\{u_{0_M}^{*}\}$ as $M \to \infty$.

PROPOSITION 3.1. *If $\tilde{u}_0^{*}$ is a cluster point in the weak $H_0^2(\Omega)$ topology of the sequence $\{u_{0_M}^{*}\}$, then $\tilde{u}_0^{*}$ is a solution of* (3.6).

Proof. Since $\tilde{u}_0^{*}$ is a cluster point of $\{u_{0_M}^{*}\}_{M=1}^{\infty}$ in the weak $H_0^2(\Omega)$ topology, there is a subsequence $\{u_{0_{M_i}}^{*}\}_{i=1}^{\infty}$ such that $u_{0_{M_i}}^{*} \to \tilde{u}_0^{*}$ weakly in $H_0^2(\Omega)$. Since $u_{0_{M_i}}^{*} \in Q_{\mathrm{ad}}^{M_i} \subset Q_{\mathrm{ad}}$ and Q_{ad} is closed and convex in $H_0^2(\Omega)$, we see that $\tilde{u}_0^{*} \in Q_{\mathrm{ad}}$. Now from the facts that $u_{0_{M_i}}^{*}$ is a solution of (3.9) and $u_{0_{M_i}}^{*} \subset Q_{\mathrm{ad}}$, we have with a solution u_0^{*} that

$$J(I^{M_i} u_0^{*}) \geqslant J(u_{0_{M_i}}^{*}) \geqslant J(u_0^{*}).$$

Since $I^{M_i} u_0^{*} \to u_0^{*}$ in $H_0^2(\Omega)$, it follows that $I^{M_i} u_0^{*} \to u_0^{*}$ in $C_0^1(\bar{\Omega})$. Thus, we see that

$$J(I^{M_i} u_0^{*}) \to J(u_0^{*})$$

as $i \to \infty$. Since weak convergence of $\{u_{0_{M_i}}^{*}\}_{i=1}^{\infty}$ in $H_0^2(\Omega)$ implies convergence in $C_0^1(\bar{\Omega})$, we have that

$$J(\tilde{u}_0^{*}) = J(u_0^{*})$$

and we conclude that $\tilde{u}_0^{*}$ is a solution of (3.6).

THEOREM 3.3. *Any sequence of solutions $\{u_{0_M}^{*}\}_{M=1}^{\infty}$ to* (3.9) *has a subsequence which converges to a solution of* (3.6).

Proof. This follows from the weak sequential compactness of Q_{ad} in $H_0^2(\Omega)$, the inclusion $Q_{\mathrm{ad}}^M \subset Q_{\mathrm{ad}}$, and Proposition 3.1.

Thus far we have determined the existence of solutions to problem (3.6) and the finite-dimensional problems (3.9). We have also obtained results concerning convergence behavior. Now we consider uniqueness and stability and show that solutions of (3.6) depend continuously upon the data. Furthermore, we investigate the uniqueness of solutions. We begin with the following observation.

Remark 3.3. Problem 3.6 is associated with the problem

$$\text{minimize } \sum_{k=1}^{N} (\varphi_k - u_{0_k})^2 \text{ subject to } \mathbf{u}_0 \in \hat{Q}_{\mathrm{ad}} \subset \mathbb{R}^N, \tag{3.10}$$

where

$\hat{Q}_{\mathrm{ad}} = \{\mathbf{u}_0 = (u_{0_1}, ..., u_{0_N}) \in \mathbb{R}^N : \exists u_0 \in Q_{\mathrm{ad}}$ with the property that $u_0(\xi_k) = u_{0_k}\}$.

Since Q_{ad} is closed, bounded, and convex in $C_0^1(\bar{\Omega})$, it follows immediately that $\hat{Q}_{\mathrm{ad}}$ is closed, bounded, and convex in $\mathbb{R}^N$. We prove only that $\hat{Q}_{\mathrm{ad}}$ is closed in $\mathbb{R}^N$.

LEMMA 3.3. *$\hat{Q}_{\mathrm{ad}}$ is closed in $\mathbb{R}^N$.*

Proof. Let $u_0^i = (u_{0_1}, ..., u_{0_N}^i) = (u_0^i(\xi_1), ..., u_0^i(\xi_N))$ belong to $\hat{Q}_{\mathrm{ad}}$ for each i and suppose that $\mathbf{u}_0^i \to \mathbf{u}_0$ in $\mathbb{R}^N$ as $i \to \infty$. Now since $u_0^i \in Q_{\mathrm{ad}}$ and Q_{ad} is weakly sequentially compact in $H_0^2(\Omega)$ (hence sequentially compact in $C_0^1(\bar{\Omega})$) there exists a subsequence in $u_0^{i_j}$ such that $u_0^{i_j} \to u_0$ in $C^0(\bar{\Omega})$ with $u_0 \in Q_{\mathrm{ad}}$. Thus, it follows that $u_0^{i_j}(x) \to u_0(x)$ for every $x \in \Omega$. In particular, we have

$$u_0^{i_j}(\xi_k) \to u_0(\xi_k)$$

for $k = 1, ..., N$. Therefore, we see that $u_{0_k} = u_0(\xi_k)$ for $k = 1, ..., N$ with $u_0 \in Q_{\mathrm{ad}}$ and $\mathbf{u}_0 \in \hat{Q}_{\mathrm{ad}}$.

Problem (3.10) is clearly a problem in a finite-dimensional Hilbert space $\mathbb{R}^N$ to find the projection of a point $\boldsymbol{\varphi} = \{\varphi_k : k = 1, ..., N\}$ onto a closed bounded convex set $\hat{Q}_{\mathrm{ad}}$ in $\mathbb{R}^N$. That this problem has a unique solution is elementary [7]. Moreover, since the solution is obtained as a projection, it follows that the solution $\mathbf{u}_0^*$ is continuous with respect to perturbation of the point $\boldsymbol{\varphi}$. In fact, if $\mathbf{u}_{01}^* = \operatorname{proj} \boldsymbol{\varphi}_1$ and $\mathbf{u}_{02}^* = \operatorname{proj} \boldsymbol{\varphi}_2$ then we have

$$|\mathbf{u}_{0_1}^* - \mathbf{u}_{0_2}^*|_{\mathbb{R}^N} \leqslant |\boldsymbol{\varphi}_1 - \boldsymbol{\varphi}_2|_{\mathbb{R}^N}, \tag{3.11}$$

where $|\boldsymbol{\varphi}| = (\sum_{k=1}^N \varphi_k^2)^{1/2}$ [6].

LEMMA 3.4. *The solution* $\mathbf{u}_0^*$ *of problem* (3.10) *is unique and continuous with respect to perturbations of the point* $\boldsymbol{\varphi}$ *and* (3.11) *holds.*

Remark 3.4. The above lemma indicates that the mapping taking $\boldsymbol{\varphi}$ to $\mathbf{u}_0^*$ is one-to-one and continuous. If Q_{ad} is infinite dimensional such as (3.5), then there may be many solutions to problem (3.6). However, we may obtain the following result.

THEOREM 3.4. *Let* S *denote the set of solutions to problem* (3.6). *The* $S \subset Q_{\text{ad}}$ *is weakly compact in* $H_0^2(\Omega)$ *and convex. If* u_0^* *and* $\hat{u}_0^*$ *belong to* S, *then for any* $x \in [\xi_1, \xi_M]$,

$$|u_0^*(x) - \hat{u}_0^*(x)\| \leqslant 2(\text{meas}\, \Omega)^{1/2} C e^{Kt_0} \Delta, \tag{3.12}$$

where Δ *is defined by* (3.14), C *is given by the right side of* (3.3), *and* $K = 1 + a + \beta$ *with* β *given in* (H2).

Proof. If $\{u_0^*\}_{j=1}^{\infty}$ is a sequence in S, then there is a subsequence $\{u_{0_{j_i}}^*\}_{i=1}^{\infty}$ that converges weakly in $H_0^2(\Omega)$ and strongly in $C_0^1(\bar{\Omega})$ to a function u_0^* in Q_{ad}. In particular, we have

$$u_{0_k}^* = u_0^{*^{j_i}}(\xi_k) \to u_0^*(\xi_k).$$

Thus, we see that $u_0^*(\xi_k) = u_{0_k}^*$ and $u_0^* \in S$. It is easy to check convexity.

For the last claim set $v(x) = u_0^*(x) - \hat{u}_0^*(x)$. Then we have for $x \in [\xi_k, \xi_{k+1}]$ that

$$|v(x)| = |v(x) - v(\xi_k)| \leqslant (\text{meas}\, \Omega)^{1/2} \|v\|_{H_0^2} |\xi_{k+1} - \xi_k|.$$

By Remark 2.1(b) we obtain for $x \in [\xi_k, \xi_{k+1}]$ that

$$|u_0^*(x) - \hat{u}_0^*(x)| \leqslant 2(\text{meas}\, \Omega)^{1/2} C\ e^{Kt_0} |(x_k, z_k) - (x_{k+1}, z_{k+1})|, \tag{3.13}$$

where C is given by the right side of the inequality (3.3), $K = 1 + a + \beta$, where β is defined in (H2), and

$$|(x_k, z_k) - (x_{k+1}, z_{k+1})| = \sqrt{(x_k - x_{k+1})^2 + (z_k - z_{k+1})^2}.$$

$$\text{Set } \Delta = \max_{1 \leqslant k \leqslant M-1} |(x_k, z_k) - (x_{k+1}, z_{k+1})|. \tag{3.14}$$

Then from (3.13) it follows that

$$|u_0^*(x) - \hat{u}^*(x)| \geqslant 2(\text{meas}\, \Omega)^{1/2} C\, e^{Kt_0}\, \Delta$$

for any $x \in [\xi_1, \xi_M]$.

Remark 3.5. We show the set of solutions is continuously dependent on the data. First we make several observations.

(i) Let S be a closed convex set in a Hilbert space H. Given an element φ in H we may define the distance from φ to S by

$$\operatorname{dist}(\varphi, S) = \min\{|\varphi - s|_H : s \in S\}.$$

This number exists [7] and is zero if and only if $\varphi \in S$.

(ii) Let S and T be two closed bounded convex sets in a Hilbert space H. We may define the distance between S and T as

$$\operatorname{dist}(S, T) = \max\left\{\max_{t \in T} \operatorname{dist}(t, S), \max_{s \in S} \operatorname{dist}(s, T)\right\}.$$

THEOREM 3.5. *Let S and $\hat{S}$ denote the solutions for problem* (3.6) *for data* $\{x_k\}_{k=1}^N$, $\{z_k\}_{k=1}^N$ *and* $\{\hat{x}_k\}_{k=1}^N$, $\{\hat{z}_k\}_{k=1}^N$, *respectively, where we view the second set as a perturbation of the first so that* $\hat{x}_k \in (x_{k-1}, x_{k+1})$ *for* $k \geqslant 2$ *and* $\hat{x}_1 \in (x_1, x_2)$. *The* $\operatorname{dist}(S, \hat{S}) \leqslant \text{Const}\, \Delta$, *where the* $\operatorname{dist}(S, \hat{S})$ *is in* $H^1(\Omega)$ *and Δ is given in* (3.14).

Proof. Let $u_0 \in S$ and $\hat{u}_0 \in \hat{S}$ and let $\xi \in [\xi_k, \xi_{k+1}]$. Then we have

$$|u_0(\xi) - \hat{u}_0(\xi)| \leqslant |u_0(\xi) - u_0(\xi_k)| + |u_0(\xi_k) - \hat{u}_0(\xi_k)| + (\hat{u}_0(\xi_k) - \hat{u}_0(\xi)|.$$

Since the derivatives of u_0 and $\hat{u}_0$ are bounded from (3.3) we see that the first and third terms are each dominated by

$$2\gamma(2\exp(-|\phi_x|\, t_0/a) - 1)(\text{meas}\,\Omega)^{1/2} C\, e^{Kt_0} \Delta,$$

where Δ is defined by (3.14). As for the second term we write

$$\begin{aligned} |u_0(\xi_k) - \hat{u}_0(\xi_k)| &\leqslant |u_0(\xi_k) - \hat{u}_0(\hat{\xi}_k)| + |\hat{u}_0(\hat{\xi}_k) - \hat{u}_0(\xi_k)| \\ &\leqslant |\zeta_k - \hat{\zeta}_k| + |\hat{u}_0(\hat{\xi}_k) - \hat{u}_0(\xi_k)|, \end{aligned}$$

where the first term may be bounded by estimates from the ordinary differential equation as Remark 2.2. The second term is bounded similarly to that above with

$$4\gamma(2\exp(-|\phi_x|\, T/a) - 1)(\text{meas}\,\Omega)^{1/2} C\, e^{kt_0}\, \Delta.$$

It follows then that

$$|u_0 - \hat{u}_0|_{C^0(\Omega)} \leqslant \text{Const}\, \Delta.$$

Moreover since $\Omega \subset \mathbb{R}^1$ we see that from a similar argument

$$|u_{0_x} - \hat{u}_{0_x}|_{C^0(\Omega)} \leqslant \text{Const } \Delta.$$

We conclude that

$$\|u_0 - \hat{u}\|_{H^1(\Omega)} \leqslant (\text{meas } \Omega)^{1/2} \text{ Const } \Delta.$$

LEMMA 3.5. *If $u_0^* \in Q_{\text{ad}}^N = S^N \cap Q_{\text{ad}}$ and there are $M \geqslant N$ observations, then the solution u_0^* is unique if and only if the matrix*

$$\begin{bmatrix} B_1(\xi_1) & B_2(\xi_1) & \cdots\cdots & B_N(\xi_1) \\ B_1(\xi_2) & B_2(\xi_2) & \cdots\cdots & B_N(\xi_2) \\ \vdots & \vdots & & \vdots \\ B_1(\xi_M) & B_2(\xi_M) & \cdots\cdots & B_N(\xi_M) \end{bmatrix} \tag{3.15}$$

has rank N. Moreover, u_0^* is continuously dependent upon $\{\xi_k\}_{k=1}^M$ and $\{u_{0_k}^*\}_{k=1}^M$.

Proof. Since $u_0 \in Q_{\text{ad}}^N$ we have $u_0(x) = \sum_{i=1}^N c_i B_i(x)$ and the solution u_0^* must satisfy

$$u_0^*(\xi_k) = \sum_{i=1}^N c_i B_i(\xi_k) = u_{0_k}^*, \qquad k = 1, ..., M.$$

This system has a unique solution if and only if (3.15) has rank N. The continuous dependence is straightforward.

THEOREM 3.6. *If $u_0^* \in Q_{\text{ad}}^N$ is the solution of* (3.9) *then u_0^* depends continuously upon the observation $(x_k\}_{k=1}^M$ and $\{z_k\}_{k=1}^M$. Further, if the matrix in* (3.15) *has rank N, then u_0^* is unique.*

Proof. That the mapping taking

$$(x_k, z_k) \to (\xi_k, \zeta_k)$$

is continuous follows from Remark 2.1(b). Thus, we have that the mapping from $(\mathbb{R}^2)^M$ into $(\mathbb{R}^2)^M$ defined by

$$(x_k, z_k)_{k=1}^M \to (\xi_k, \zeta_k)_{k=1}^M$$

is continuous. The mapping to u_0^* may be viewed as the composition

$$(x_k, z_k)_{k=1}^M \to (\xi_k, \zeta_k)_{k=1}^M \to (\xi_k, u_{0_k}^*)_{k=1}^M \to u_0^*.$$

The continuity and injectivity of the last two mappings follow from Lemma 3.4 (continuity of the projection onto a closed convex set) and Lemma 3.5, respectively.

Remark 3.6. The above discussion has demonstrated that for the infinite-dimensional case, although the solution of problem (3.6) is not unique, the set of solutions S is closed convex with diameter determined by the space of observations. Moreover, the set S is continuous with respect to perturbations of the observations. In the discrete case we have determined necessary and sufficient conditions for uniqueness and continuity of the solutions upon observations.

4. A Numerical Example

In the section we describe a numerical experiment for the problem

$$\text{minimize } J(u_0) = \sum_{k=1}^{M} (u_0(\xi_k) - \zeta_k)^2 \text{ subject to } U_0 \in Q_{\text{ad}}. \tag{4.1}$$

This example is solely for illustration of certain aspects of the theory. We make no claim that the techniques used are optimal. The observations $\{\xi_k\}_{k=1}^M$ and $\{\zeta_k\}_{k=1}^M$ are obtained from observed data $\{x_k\}_{k=1}^M$ and $\{z_k\}_{k=1}^M$ at t_0 by solving the initial value problem

$$\begin{aligned} \frac{dx}{dt} &= z, & x(t_0) &= x_k \\ \frac{dz}{dt} &= -az + \phi(x, t), & z(t_0) &= z_k \end{aligned} \tag{4.2}$$

and setting

$$\xi_k = x(0; x_k, z_k) \quad \text{and} \quad \zeta_k = z(0; x_k, z_k).$$

For our sample problem we make the assumptions

$$a = 1.0, \qquad t_0 = 0.1, \qquad \phi(x, t) = 1.0 \tag{A1}$$

$$u_0(x) = (\pi/2 + \arctan x)/2. \tag{A2}$$

To generate data we solve the problem

$$\begin{aligned} \frac{dx}{dt} &= z, & x(0) &= \hat{\xi}_k \\ \frac{dz}{dt} &= -z + 1, & z(0) &= u_0(\hat{\xi}_k) \end{aligned} \tag{4.3}$$

analytically forward to time $t=t_0$ from specified $\{\hat{\xi}_k\}_{k=1}^M$ to obtain

$$\begin{aligned} x(t_0;\hat{\xi}_k, u_0(\hat{\xi}_k)) &= \hat{\xi}_k + t_0 + (u_0(\hat{\xi}_k)-1)(1-e^{-t_0}) \\ z(t_0;\hat{\xi}_k, u_0(\hat{\xi}_k)) &= 1 + (u_0(\hat{\xi}_k)-1)\, e^{-t_0} \end{aligned} \tag{4.4}$$

for each $k=1, ..., M$. These values are then perturbed by adding or subtracting 0.001 to obtain the data $\{x_k\}_{k=1}^M$, $\{z_k\}_{k=1}^M$ for the test problem. The function u_0 is smooth monotone increasing and does not give rise to a shock within the time interval $(0, t_0)$. We then solve (4.2) backward from $t=t_0$ to $t=0$ by using the classical fourth-order Runge–Kutta scheme [4]. In this way we obtain $\{\xi_k\}_{k=1}^M$ and $\{\zeta_k\}_{k=1}^M$.

To solve (4.1) numerically we consider in particular finite-dimensional cases

$$Q_{\text{ad}} = Q_{\text{ad}}^N = \left\{u_0 : u_0(x) = \sum_{i=1}^N c_i B_i(x)\right\} \tag{A3}$$

$$N = M. \tag{A4}$$

The functions B_i are obtained from the reference element [9, 10]

$$B_{\text{ref}}(t) = \begin{cases} (t+2)^3, & -2 \leqslant t \leqslant -1 \\ 1+3(t+1)+3(t+1)^2-3(t+1)^3, & -1 \leqslant t \leqslant 0 \\ 1+3(1-t)+3(1-t)^2-3(1-t)^3, & 0 \leqslant t \leqslant 1 \\ (2-t)^3, & 1 \leqslant t \leqslant 2 \\ 0, & \text{otherwise,} \end{cases}$$

where in general for $x_i = (i-1)/(N-1)$, $i=1, ..., N$, we set

$$B_i(x) = B_{\text{ref}}((x-x_i)/(N-1)).$$

Assumption (A4) is simply a matter of convenience and is consistent with Lemma 3.5. In this example Q_{ad} is essentially unconstrained and problem (4.1) is solved as an unconstrained problem. This is due to the fact that a shock does not develop from the test data u_0 from (A2).

To solve (4.1) we use the steepest descent method [7]. In the case that Q_{ad} is given as in (A2)–(A4), we have for $c=(c_1, ..., c_N)$ that

$$j(c) = J(u_0) = \sum_{k=1}^N \left(\zeta_k - \sum_{i=1}^N c_i B_i(\xi_k)\right)^2. \tag{4.5}$$

The partial derivatives are given by

$$\frac{\partial j(x)}{\partial c_l} = -z \sum_{k=1}^N \left(\zeta_k - \sum_{i=1}^N c_i B_i(\xi_k)\right) B_l(\xi_k) \tag{4.6}$$

with the derivative

$$D^j(c) = \left(\frac{\partial j(c)}{\partial c_l} : l = 1, ..., N \right).$$

Given the coefficient $c^{(0)}$, the steepest descent method defines the next coefficient $c^{(1)}$ by

$$c^{(1)} = c^{(0)} - \alpha D^j(c^{(0)}), \tag{4.7}$$

where $\alpha > 0$ is a step-size parameter and is chosen to minimize

$$j(c^{(0)} - \alpha D^j(c^{(0)})).$$

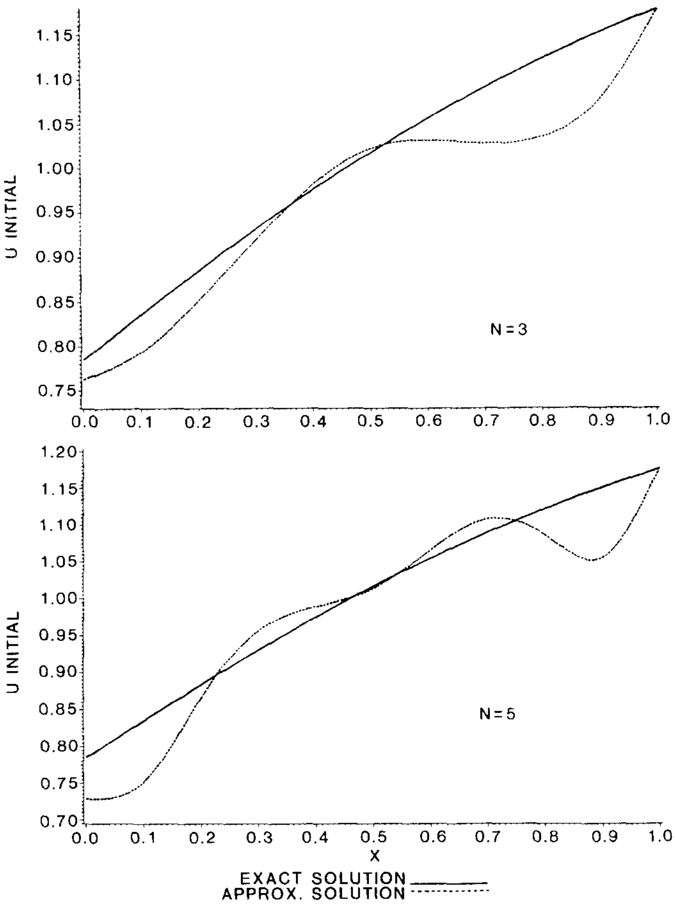

FIGURE 1

Since $j(x)$ is a quadratic functional, we may explicitly calculate α to be given by

$$\alpha = \frac{\sum_{k=1}^{N} [(u_0(\xi_k) - \zeta_k) \sum_{i=1}^{N} (\partial j)/(\partial c_i)(c^{(0)}) B_i(\xi_k)]}{\sum_{i=1}^{N} (\sum_{i=1}^{N} (\partial j)/(\partial c_i)(c_0) B_i(\xi_k))^2}. \tag{4.8}$$

The procedure is then repeated with $c^{(0)}$ replaced by $c^{(1)}$.

Using data from points $\{\xi_k\}_{k=1}^{N}$ given by $\xi_k = k/(N-1)$, $k = 0, 1, ..., N-1$, basis functions B_i are defined on the mesh ξ_k and we give the results in the following table where we record the $L2$ Relative Error. That is, we use

$$L2 \text{ Rel Err} = \frac{1}{11} \sum_{k=0}^{10} \left[\frac{u_0(s_k) - \sum_{i=1} c_i B_i(s_k)}{u_0(s_k)}\right]^2,$$

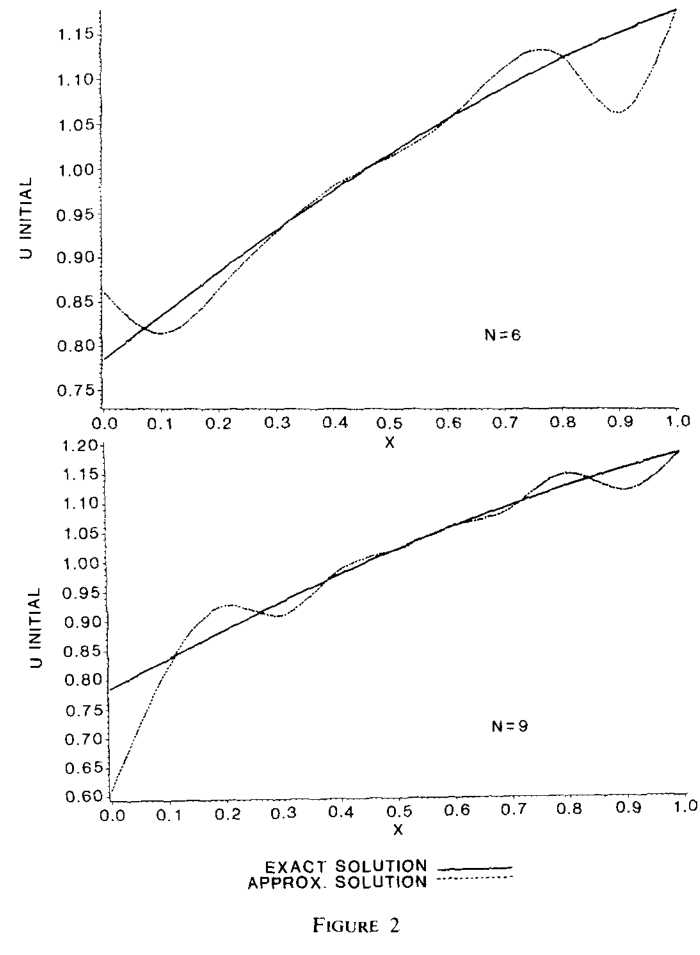

FIGURE 2

where $s_k = k/10$.

N	$L2$ Rel. Error
3	0.03127
5	0.02176
6	0.00086
9	0.00049

In Figs. 1 and 2, we graphically depict the estimated u_0 along with the exact solution u_0 given in (A2).

REFERENCES

1. R. A. ADAMS, "Sobolev Spaces," Academic Press, New York, 1975.
2. E. CODDINGTON AND N. LEVINSON, "Theory of Ordinary Differential Equations," McGraw-Hill, New York, 1955.
3. J. DIEUDONNÉ, "Foundations of Modern Analysis," Academic Press, New York, 1960.
4. C. W. GEAR, "Numerical Initial Value Problems in Ordinary Differential Equations," Prentice-Hall, Englewood Cliffs, NJ, 1971.
5. J. R. HOLTON, "An Introduction to Dynamic Meteorology," Academic Press, New York, 1972.
6. D. KINDERLEHRER AND G. STAMPACCHIA," "An Introduction to Variational Inequalities and Their Applications," Academic Press, New York, 1980.
7. D. G. LUENBERGER, "Optimization by Vector Space Methods," Wiley, New York, 1969.
8. P. MOREL, An overview of meteorological data assimilation, *in* "Seminar on Data Assimilation Methods, 1980," pp. 1–18, European Centre for Medium Range Weather Forecasts.
9. P. M. PRENTER, "Splines and Variational Methods," Wiley, New York, 1975.
10. M. SCHUTZ, "Spline Analysis," Prentice-Hall, Englewood Cliffs, NJ. 1973.
11. J. SMOLLER, "Shock Waves, and Reaction-Diffusion Equations," Springer-Verlag, New York, 1980.
12. L. W. WHITE, Estimation of parameters in nonlinear flow equations with discontinuous solutions, submitted for publication.